

# Prediction of Soil Sand and Clay Contents via Visible and Near-Infrared (Vis-NIR) Spectroscopy

Zeynal Tümsavaş<sup>a</sup>, Yücel Tekin<sup>b, 1</sup>, Yahya Ulusoy<sup>b</sup>, Abdul M. Mouazen<sup>c</sup>

<sup>a</sup>Uludag University, Agricultural Faculty, Bursa, Turkey

<sup>b</sup>Uludag University, Vocational School of Technical Sciences, Bursa, Turkey

<sup>c</sup>Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium.

**Abstract.** Visible and near infrared (Vis-NIR) spectroscopy is a non-destructive analytical method that can be used to complement, enhance or potentially replace conventional methods of soil analysis. The aim of this research was to predict the particle size distribution (PSD) of soils using a Vis-NIR spectrophotometry in one irrigate field having a vertisol clay texture in the Karacabey district of Bursa Province, Turkey. A total of 86 soil samples collected from the study area were subjected to optical scanning in the laboratory with a portable, fiber-type Vis-NIR spectrophotometer (AgroSpec, tec5 Technology for Spectroscopy, Germany). Before the partial least square regression (PLSR) analysis, the entire reflectance spectra were randomly split into calibration (80%) and validation (20%) sets. A leave-one-out cross-validation PLSR analysis was carried out using the calibration set with Unscrambler® software, whereas the model prediction ability was tested using the validation (prediction) set. Models developed were used to predict sand and clay content using on-line collected spectra from the field. Results showed an “excellent” laboratory prediction performance for both sand ( $R^2 = 0.81$ , RMSEP = 3.84% and RPD = 2.32 in cross-validation;  $R^2 = 0.90$ , RMSEP = 2.91% and RPD = 2.99 in the prediction set) and clay ( $R^2 = 0.86$ , RMSEP = 3.4% and RPD = 2.66 in cross validation;  $R^2 = 0.92$ , RMSEP = 2.67% and RPD = 3.14 in the prediction set). Modelling of silt did not result in any meaningful correlations. Less accurate on-line predictions were recorded compared to the laboratory results, although the on-line predictions were very good (RPD = 2.24-2.31). On-line predicted maps showed reasonable spatial similarity to corresponding laboratory measured maps. This study proved that soil sand and clay content can be successfully measured and mapped using Vis-NIR spectroscopy under both laboratory and on-line scanning conditions.

**Keywords.** PLS regression analysis, sand, clay, Vis-NIR spectroscopy

## 1. Introduction

Soil is one of the most important natural resources on the earth, and it has numerous characteristics resulting from the effects of climate on parent material in a specific topography and from biotic activities over a certain period of time. For sustainable land management, it is necessary to understand the characteristics of soils, their functions and spatial and temporal changes [1, 2]. It is possible to determine the physical, chemical and biological properties of the soil in order to reveal its potential as well as its limitations for agricultural and non-agricultural land use.

---

<sup>1</sup> Corresponding Author, Gorukle Campus, 16059, Bursa, Turkey. E-mail: ytekin@uludag.edu.tr

Soil texture, represented by the distribution of the size of mineral particles in the soil (particle size distribution - PSD) is one of the most basic static physical properties of soils. Soil texture can exhibit significant spatial variation within a land area. Simply defined, soil texture is the proportional distribution of sand, silt and clay particles in a soil mass. Many factors are affected by soil texture, including plant growth and yield, the infiltration of water into the soil and its storage, retention and transport, the availability and absorption of plant nutrients, living organisms in the soil, soil quality and productivity, soil temperature, structure and compaction levels, tillage, irrigation, and the effectiveness of fertilizers. Thus, the concept of soil texture has a much greater importance beyond its simple definition and plays a key role in agricultural production. Measuring the spatial variation in soil texture can be of great benefits to site specific land management.

Many soil analysis laboratories generally use the traditional hydrometer and pipette methods to determine PSD. Some errors may result from the use of these methods in routine laboratory analysis. In their particle size analysis study, Klein et al. [3] observed that the highest numbers of particle size fraction classification errors were found in those made for silt and clay. In addition, it has been reported that the inconsistency of obtained clay fraction values was due to the difficulty involved in laboratory analysis of the dispersion of clay particles [4]. The traditional analysis methods for PSD used in soil laboratories are expensive, labour-intensive and require more preparation, especially when working with a large number of soil samples. Moreover, since the analysis process takes considerable time with these methods, they are not practical for application in precision agriculture, high resolution soil mapping or in soil surveys carried out over large areas. Therefore, alternative methods need to be developed for the determination of the proportions of soil mineral fractions. In recent years, various electromagnetic radiation techniques have been used in the prediction of a number of soil properties, and it has been reported that visible (Vis), near infrared (NIR) and mid infrared (MIR) spectrophotometry, nuclear magnetic resonance (NMR) and mass spectrophotometry (MS) could complement traditional laboratory soil analysis methods as alternative techniques [5]. The main components of the soil such as its clay content and mineralogy, the amounts of organic matter and iron oxides and soil moisture, texture and particle size directly affect the spectral behavior of the soil [6, 7]. Vis-NIR reflectance spectroscopy shows promise as an alternative method, making it possible to measure many soil properties at the same time. When compared to traditional laboratory analyses, Vis-NIR reflectance spectroscopy results have been successful to a degree in the determination of soil physical properties like bulk density, soil texture and structure [8, 9, 10], but more successful in the determination of soil content of total nitrogen different forms of carbon, cation exchange capacity (CEC) and soil chemical properties such as pH and P [11, 12, 13, 14, 15, 16, 17]. The Vis-NIR reflectance spectroscopy method is readily adaptable for both in-laboratory and in-situ measurements, and requires very little or no soil preparation using chemical reagents. It is a quick and inexpensive method that does not destroy the sample [11, 18, 19]. Laboratory and in situ applications are possible with the Vis-NIR spectroscopy for the analyses of soil. However, no study on on-line measurement of soil texture fractions can be found in the literature, although measurements under laboratory conditions were performed by many researchers using the Vis-NIR spectroscopy.

The aim of this study was to determine the potential for use of Vis-NIR spectrophotometry in the prediction of the PSD (sand, silt and clay) of the soil under non-mobile (laboratory) and on-line (mobile) measurement conditions.

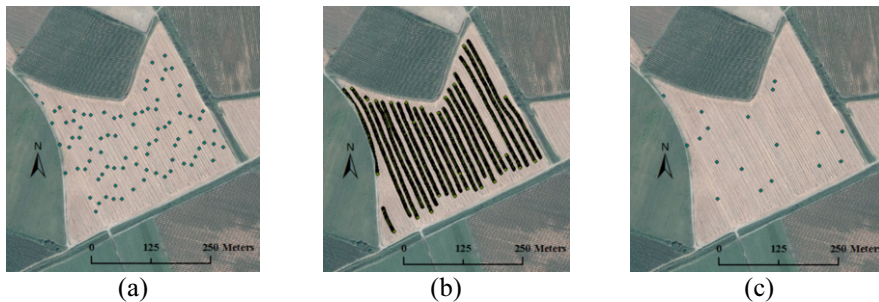
## 2. Materials and Methods

### 2.1. On-line soil sensor

The on-line sensor system consists of a subsoiler with an optical probe mounted on the rear of a subsoiler, installed on a tractor-pulled framework [20]. The system was produced at Uludağ University in Bursa, Turkey, using the same design patented by Mouazen [20]. A mobile, fiber type, AgroSpec Vis-NIR spectrophotometer (tec5 Technology for Spectroscopy, Germany) was used to measure soil spectra. A differential global positioning system (DGPS) (EZ-Guide 250 Trimble, USA) was used to record with submeter accuracy the on-line measured spectral positions. In the spectral measurement system, AgroSpec software (tec5 Technology for Spectroscopy, Germany) was employed for the simultaneous collection of the spectral and GPS data.

### 2.2. Study area and on-line measurement

This study was carried out on a 10.06-ha area of agricultural land in the district of Karacabey in Bursa Province, Turkey. A total of 86 soil samples were collected from the bottom of trench opened by the subsoiler in the study area during the on-line measurement. While the tractor was moving at a speed of approximately  $3 \text{ km/h}^{-1}$ , the raw reflectance values were collected from the bottom of the trench opened along straight lines parallel to each other with 10 m interval. For the purpose of validation, at approximately 20 m intervals along these lines, soil samples were collected from the bottom of the furrows, and then put in nylon packets and numbered for laboratory analysis. The locations of the points where the samples were taken were recorded via DGPS. The sampling lines and sampling points are given in Figure 1.



**Figure 1.** (a) Soil sampling points; (b) on-line soil measurement transects; (c) validation points (Ulusoy et al., 2016)

Each of the 86 soil samples taken from the study area was divided into two parts. One part of the soil sample was used for the laboratory reference PSD measurement, while the second part was used for optical scanning in the laboratory. The PSD was determined using the Bouyoucos hydrometer method [21]. According to the results of the PSD analysis, the texture classes of the soil were determined using the classification system of the United States Department of Agriculture (Table 1) [22].

**Table 1.** Particle size distribution (PSD) of study area soil

Area (ha)	Crop	Texture class	Sand (%)	Silt (%)	Clay (%)
10	wheat	Clay	26.6	30.4	43.0

2.3. Optical measurement in the laboratory

The soil samples were scanned in the laboratory using the same Vis-NIR spectrophotometer (AgroSpec, tec5 Technology for Spectroscopy, Germany) employed during on-line measurements in the field. After collection, root residues, stubble and gravel were removed and each soil sample was then thoroughly mixed before being scanned with the Vis-NIR spectrophotometer. The soil sample was then distributed into three plastic cup, each having a depth of 2.5 cm and a diameter of 4.7 cm, and the soil was carefully levelled in order to ensure a smooth scanning surface for the soil in the containers [23]. Before beginning the soil sample scanning process, a 100% white reference was scanned by the spectrophotometer and this was repeated every 30 min. Each of the soil samples in the containers was scanned 10 times and the average of the readings was considered for further analyses. By taking the average of the three reflection values collected from the three containers, the final reflection value to be used in determining the properties of the soil sample was thus obtained.

2.4. Modelling

Before the partial least square regression (PLSR) analysis, the entire sample set (86 samples) were randomly split into calibration (80%) and validation (20%) sets. A leave-one-out cross-validation PLSR analysis was carried out using the calibration set with Unscrambler® software, whereas the model prediction ability was tested using the validation (prediction) set. The calibration and prediction sets of the sand and clay measured in the laboratory and on-line and the sample statistics are given in Table 2. Calibration models were developed for the three texture fractions, namely, sand, silt and clay. Calibration models were validated using the spectra of the validation set, scanned in the laboratory under stationary conditions and in the field under mobile on-line conditions. These calibration models were used to predict sand and clay content using on-line collected spectra from the field. Since no correlation for silt content was possible with the PLSR cross-validation, no model was established and on-line prediction of silt was excluded.

The evaluation of model performance was made by examining  $R^2$  value, the root mean square error of prediction (RMSEP) and the residual prediction deviation (RPD).

**Table 2.** Sample statistics for calibration and prediction sets of laboratory and on-line measured sand and clay

		Sample number	Min (%)	Max (%)	Mean (%)	SD (%)
All samples	Sand	86	12.4	55.8	24.57	9.03
	Clay	86	23.3	62.3	45.85	9.08
Cross-validation set	Sand	68	12.4	55.8	24.20	8.95
	Clay	68	23.3	62.3	46.08	9.06
Laboratory prediction set	Sand	18	13.93	49.9	25.96	9.45
	Clay	18	24.3	60.3	44.99	9.37
On-line prediction set	Sand	18	8.88	49.7	26.10	9.78
	Clay	18	24.67	59.9	44.29	9.42

### 2.5. Development of sand and clay maps

Five categories of sand and clay maps were developed. These included: (1) a laboratory analysis map based on the laboratory measurements of 86 points, (2) a map of the laboratory prediction set (18 points), (3) an on-line prediction map based on 18 verification points, (4) a map of laboratory reference values for 18 points and (5) an on-line Vis-NIR map based on all 8486 predicted measurement points. The kriging method after semi-variogram analysis was employed to draw the on-line map showing all the predicted points, while for the other maps, the inverse distance weighing (IDW) interpolation method was used. All maps were generated via ArcGIS 10 (ESRI, USA) software [24].

## 3. Results and Discussion

### 3.1. Calibration and prediction model performance

The performance of the PLSR model in cross-validation and the laboratory and on-line predictions for the sand and clay content of the study field is shown in Table 3. According to the classification of RPD values suggested by Viscarra-Rossel et al. [11], the performance of the sand model in the cross-validation is classified as “very good” ( $R^2 = 0.81$ , RMSEP = 3.84 and RPD = 2.32).

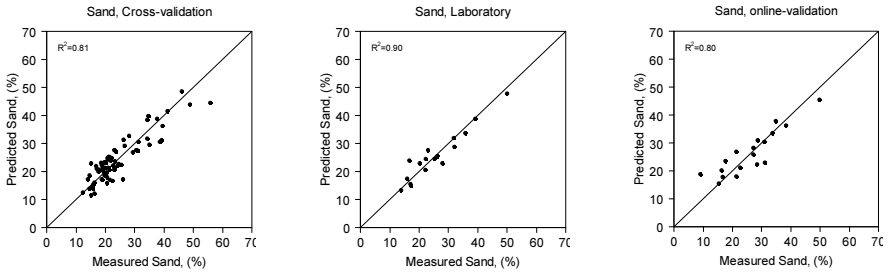
**Table 3.** Sand and clay model performance in cross-validation, laboratory and on-line predictions

		$R^2$	RMSEP (%)	RPD	Intercept	Slope
Cross-validation set	Sand	0.81	12.4	2.32	4.27	0.73
	Clay	0.85	23.3	2.66	6.55	0.92
Laboratory prediction	Sand	0.90	12.4	2.99	3.04	0.87
	Clay	0.91	23.3	3.14	6.48	0.85
Cross-validation	Sand	0.80	8.88	2.24	7.17	0.73
	Clay	0.82	24.67	2.31	11.4	0.72

RMSEP: Root mean square error of prediction  
RPD: Residual prediction deviation

The performance of Vis-NIR prediction set model for sand content under on-line measurement conditions are not as good as those under laboratory measurement conditions. According to the classification of RPD values suggested by Viscarra-Rossel et al. [11], the on-line prediction model, with RPD of 2.24, is classified as “very good” (RPD 2.0 – 2.5), while the laboratory prediction model, with RPD of 2.99, is classified as “excellent” (RPD > 2.5). Scatter plots of the cross-validation set, laboratory and on-line validation results of sand content are given in Figure 2.

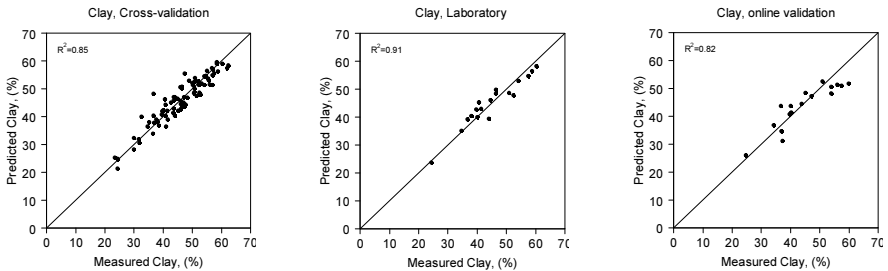
The performance of Vis-NIR prediction set model of for sand content under on-line measurement conditions are not as good as those under laboratory measurement conditions. According to the classification of RPD values suggested by Viscarra-Rossel et al. [11], the on-line prediction model, with RPD of 2.24, is classified as “very good” (RPD 2.0 – 2.5), while the laboratory prediction model, with RPD of 2.99, is classified as “excellent” (RPD > 2.5). Scatter plots of the cross-validation set, laboratory and on-line validation results of sand content are given in Figure 2.



**Figure 2.** Scatter plot of predicted versus laboratory-measured sand of (a) the cross-validation set, (b) the prediction set for 18 laboratory-scanned samples, and (c) the prediction set for 18 on-line-scanned samples

Just like with the sand, the performance of Vis-NIR models for the prediction of clay content under on-line measurement conditions are not as good as those under laboratory measurement conditions (Table 3). According to the RPD classification system of Viscarra-Rossel et al. [11], the on-line model, with an RPD of 2.31, is classified as “very good”, while the prediction model based on measurements made in the laboratory, with an RPD of 3.14, is classified as “excellent” (RPD value > 2.5).

The clay content measured in the laboratory versus the predicted clay content in the prediction sets of soil reflection spectra scanned in the laboratory and scanned on-line, and in the cross-validation set can be seen in the scatter plots in Figure 3.

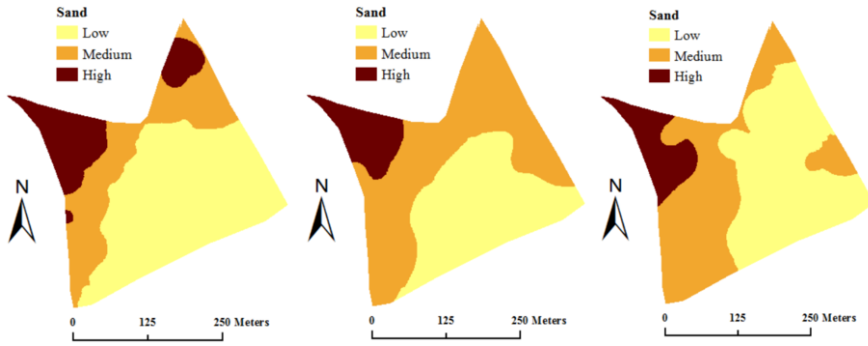


**Figure 3.** Scatter plot of predicted versus laboratory-measured clay of (a) the cross-validation set, (b) the prediction set for 18 laboratory-scanned samples and (c) the prediction set for 18 on-line-scanned samples

### 3.2. Mapping

#### 3.2.1. Comparison of laboratory and Vis-NIR maps of soil sand and clay content

The maps of the laboratory-measured, laboratory Vis NIR-predicted and on-line Vis-NIR-predicted data for sand content using the prediction set of 18 soil samples show acceptable spatial similarity when compared. In addition, the low and high sand areas of the field can be clearly distinguished (Figure 4).



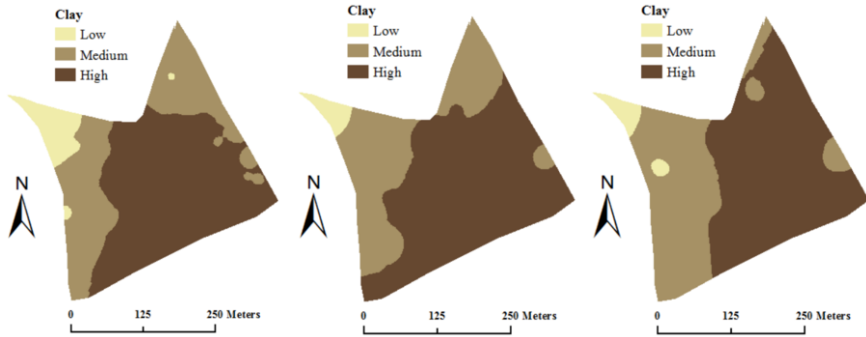
**Figure 4.** Comparison of (a) laboratory measured (86 points), (b) laboratory visible and near infrared (Vis-NIR) predicted and (c) on-line Vis-NIR predicted maps of sand content (based on the 18 samples of the prediction set)

With minor differences, in general, all the maps show the central-southern part of the study area as having a lower sand content than other parts (Figure 4). The similarity between the laboratory reference (18 samples) and the whole-point (8486 points) sand content maps can be seen in Figure 5. Compared to other parts, the central-southern part of the field on these two maps also exhibits a spatial distribution similar to that seen on the three maps in Figure 4, indicating that it has a lower sand content.



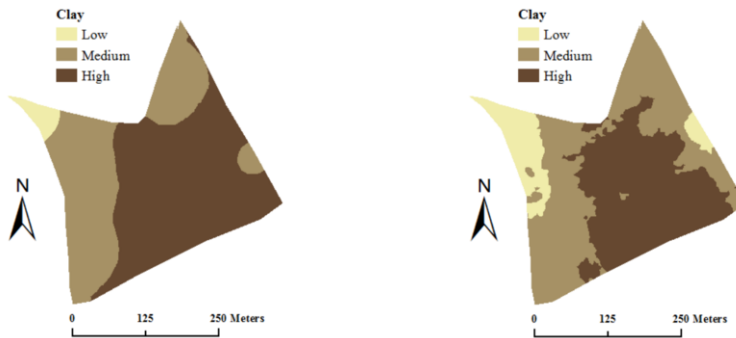
**Figure 5.** Comparison between (a) the 18 laboratory reference point and (b) the full-point on-line visible and near infrared (Vis-NIR) predicted maps for sand content.

In contrast to the spatial distribution maps of the sand content in the study site (Figure 4), the maps prepared for the clay content using the prediction set of 18 soil samples, based on the laboratory-measured, laboratory Vis-NIR-predicted and on-line Vis-NIR-predicted data, showed very close spatial similarity when compared, and the low, medium and high clay content zones can also be clearly distinguished (Figure 6).



**Figure 6.** Comparison of (a) laboratory-measured, (b) laboratory visible and near infrared (Vis-NIR) predicted and (c) on-line Vis-NIR-predicted maps of clay content (based on the 18 samples of the prediction set)

As in the spatial distribution maps of the sand content (Figure 4), there are also minor differences seen in all the spatial distribution maps of the of clay content. These maps indicated that the central-southern part of the field has a higher clay content than the other parts (Figure 6). The similarity between the laboratory reference (18 samples) and the whole-point (8486 points) maps of clay content can be seen in Figure 7. These two maps show a very close spatial distribution similarity to the three maps in Figure 6, indicating that the central-southern parts of the field has a higher clay content than the other parts. The opposite spatial distribution of clay compared to that of sand indicates the good model performance in predictions of clay and sand content.



**Figure 7.** Comparison between (a) the laboratory reference and (b) the full-point on-line visible and near infrared (Vis-NIR) predicted maps for clay content

The high sampling resolution map obtained via the on-site measuring soil sensor has provided detailed information on the spatial distribution of the sand and clay content of the soil. This information is very useful for sustainable and precision agricultural applications and for land, soil and plant management. In addition, it can give direction to agronomists on the economic use of resources and contribute to the planning of plant production strategies.



#### 4. Conclusions

This study evaluated the potential of visible and near infrared (Vis-NIR) spectrophotometry in the determination of the sand and clay content in clay-textured field soil under the semi-humid climate conditions in Turkey. The following conclusions have been determined according to the results obtained under laboratory and in-situ (on-line) field measurement conditions:

1. Vis-NIR spectrophotometry can be used successfully to determine and map the sand and clay in clay-textured soils in a semi-humid climate region.
2. Vis-NIR scanning under laboratory conditions as opposed to on-line Vis-NIR measurement conditions can be expected to provide better measurement accuracy.
3. Vis-NIR-predicted sand and clay maps and the equivalent laboratory-measured maps showed significant similarities. However, when the map developed using a limited number of points (18 samples) as compared with the full-point map, the full-point map showed more detail and displayed slightly different spatial distribution patterns.

As future work, Vis-NIR spectrophotometry application could be extended to provide on-line measurement sand and clay content in soils, to work that will link crop yield with plant characteristics.

#### References

- [1] McBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S., Shatar, T.M., An overview of pedometric techniques for use in soil survey, *Geoderma* **97** (2000), 293-327.
- [2] Florinsky, I.V., Eilers, R.G., Manning, G.R., Fuller, L.G., Prediction of soil properties by digital terrain modeling, *Environmental Modelling and Software* **17** (2002), 295-311.
- [3] Klein, V.A., Madalosso, T., Reichert, J.M., Suzuki, L.E.A.S., Veiga, M., Albuquerque, J.A., Pauletto, E.A., Methodologies for quality control of soil particle size analysis, *Ciência Rural* **43** (2013), 850-853.
- [4] Suzuki, L.E.A.S., Reichert, J.M., Albuquerque, J.A., Reinert, D.J., Kaiser, D., Dispersion and flocculation of Vertisols, Alfisols and Oxisols in Southern Brazil, *Geoderma Regional* **5** (2015), 64-70.
- [5] Janik, L.J., Merry, R.H., Skjemstad, J.O., Can mid infrared diffuse reflectance analysis replace soil extractions, *Australian Journal of Experimental Agriculture* **38** (1998), 681-696.
- [6] Dalmolin, R.S.D., Gonçalves, C.N., Klamt, E., Dick, D.P., Relationship between the soil constituents and its spectral behavior, *Ciência Rural* **35** (2005), 481-489.
- [7] Mouazen, A.M., Karoui, R., Deckers, J., De Baerdemaeker, J., Ramon, H., Potential of visible and near-infrared spectroscopy to derive colour groups utilising the Munsell soil colour charts, *Biosystems Engineering* **97** (2007), 131-143.
- [8] Cécillon, L., Barthes, B.G., Gomez, C., Ertlen, D., Genot, V., Hedde, M., Stevens, A., Brun, J.J., Assessment and monitoring of soil quality using near-infrared reflectance spectroscopy (NIRS), *European Journal of Soil Science* **60** (2009), 770-784.
- [9] Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.M., McBratney, A., Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy, *Trends Analytical Chemistry* **29** (2010), 1073-1081.
- [10] Gholizadeh, A., Amin, M.S.M., Boruovka, L., Saberioon, M.M., Models for estimating the physical properties of paddy soil using visible and near infrared reflectance spectroscopy, *Journal of Applied Spectroscopy* **81** (2014), 534-540.
- [11] Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties, *Geoderma* **131** (2006), 59-75.

- [12] Tekin, Y., Tümsavaş, Z., Mouazen, A.M., Effect of Moisture Content on Prediction of Organic Carbon and pH Using Visible and Near-Infrared Spectroscopy, *Soil Science Society of America Journal* **76** (2012), 188-198.
- [13] Tümsavaş, Z., Tekin, Y., Mouazen, A.M., Effect of moisture content on the prediction of cation exchange capacity using visible and near infrared spectroscopy, *Journal of Food Agriculture and Environment* **11** (2013), 760-764.
- [14] Tekin, Y., Tümsavaş, Z., Mouazen, A.M., Comparing the artificial neural network with partial least squares for prediction of soil organic carbon and pH at different moisture content levels using visible and near-infrared spectroscopy, *Revista Brasileira de Ciência Solo* **38** (2014), 1794-1804.
- [15] Ji, W., Viscarra Rossel, R.A., Shi, Z., Improved estimates of organic carbon using proximally sensed vis-NIR spectra corrected by piecewise direct standardization, *European Journal of Soil Science* **66** (2015), 670-678.
- [16] Kuang, B.Y., Tekin, Y., Mouazen, A.M., Comparison between artificial neural network and partial least squares for on-line visible and near infrared spectroscopy measurement of soil organic carbon, pH and clay content, *Soil and Tillage Research* **146** (2015), 243-252.
- [17] Ulusoy, Y., Tekin, Y., Tümsavaş, Z., Mouazen, A.M., Prediction of soil cation exchange capacity using visible and near infrared spectroscopy, *Biosystems Engineering* **152** (2016), 79-93.
- [18] McCarty, G.W., Reeves III, J.B., Reeves, V.B., Follett, R.F., Kimble, J.M., Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement, *Soil Science Society of America Journal* **66** (2002), 640-646.
- [19] Vasques, G.M., Grunwald, S., Sickman, J.O., Comparison of multivariate methods for inferential modeling of soil carbon using visible/ near-infrared spectra, *Geoderma* **146** (2008), 14-25.
- [20] Mouazen, A.M., 2006. Soil Survey Device. International publication published under the patent cooperation treaty (PCT). World Intellectual Property Organization, International Bureau. International Publication Number: WO2006/015463, PCT/BE2005/000129, IPC: G01N21/00, G01N21/00.
- [21] Gee, G.W., Bauder, J.W., *Methods of Soil Analysis*, America Society of Agronomy Inc., Soil Science Society of America Inc., Madison, WI, USA, 1986.
- [22] Soil Survey Staff, *Keys to Soil Taxonomy*, United States Department of Agriculture, Natural Resources Conservation Service, Washington, DC. 1998.
- [23] Mouazen, A. M., De Baerdemaeker, J., Ramon, H., Towards development of on-line soil moisture content sensor using a fibre-type NIR spectrophotometer, *Soil and Tillage Research* **80** (2005), 171-183.
- [24] ArcGis help. Available online: <http://resources.arcgis.com/en/help/> (accessed on 03 March 2014).